MATH3611 / MATH5705

Chapter 1: Sets and cardinality

Section 1

Origins of set theory

Set theory and the foundations of mathematics

Georg Cantor, founder of set theory

Question: Can a set be a member of itself?

Intuitively, a set is any "collection of things". For a member x of a set S, we use the notation

$$x \in S.$$

However, a naive understanding of sets as arbitrary collections quickly leads us into paradoxes. We might ask whether sets can be members of themselves: can we have

$$S \in S?$$

It seems the answer should be *yes*: if we consider the "set of all sets", then by assumption this is a set, so in particular it is a member of itself.

Russell's paradox: Let S be the set of all sets which are *not* members of themselves:

$$S = \{T : T \text{ is a set and } T \notin T\}.$$

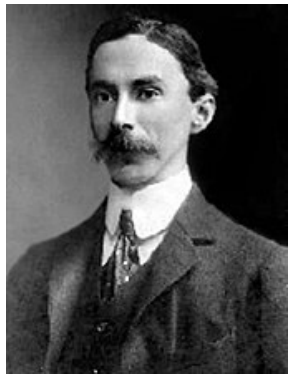
Is S a member of itself?

Question:

Is S actually a member of itself?

- If *yes*, then $S \in S$. So S *does not* satisfy the defining condition for membership in S, and therefore $S \notin S$.
- If *no*, then $S \notin S$. So S *does* satisfy the defining condition for membership in S, and therefore $S \in S$.

So either way we get a contradiction.



Bertrand Russell

Axiomatic set theory


Ernst Zermelo and Abraham Frankel

The standard axiomatization of set theory is called *Zermelo-Frankel (ZF)* set theory.

Russell's paradox shows that a naive definition of sets can lead to problems. In particular, it does not make sense to talk about the "set of all sets" or similarly large sets.
Instead, a more rigorous approach is given by *axiomatic set theory*. To study axiomatic set theory properly would take an entire course, so instead we will just briefly mention some main points which are important for this course in analysis.

Section 2
Zermelo-Frankel set theory and the Axiom of Choice

The standard axiomatization of set theory is called *Zermelo-Frankel (ZF)* set theory.

Ernst Zermelo and Abraham Frankel

We will not study ZF in detail, but the axioms assume/imply the existence of some basic sets (the empty set and an infinite set), and also give some ways of constructing new sets from old, including:

Constructing sets: *union*, *subset*, *power set*

- Taking unions: If $S = \{T_i\}_{i \in I}$ is a set of sets, then the *union*

$$\bigcup_{i \in I} T_i = \{x : \exists i \in I \text{ such that } x \in T_i\}$$

  (the set whose members are elements which belong to at least one of the T_i) is a set.
- *Subsets* with a specified condition: If S is a set and ϕ is a "condition" on elements, then

$$\{x \in S : \phi(x)\}$$

  (the subset of S consisting of those elements in S for which the condition ϕ holds) is a set.
- Power set: If S is a set, then the *power set*

$$\{T : T \subseteq S\}$$

  (the set of all subsets of S) is a set.

Using the ZF axioms one can define numbers, develop arithmetic, and describe lots of other mathematical notions.

Cartesian product

The ZF axioms allow one to construct Cartesian products of indexed sets. If A and B are two sets, then the Cartesian product $A \times B$ is the set of all pairs, or "2-tuples", (a, b) such that $a \in A$ and $b \in B$. More generally, if $\{S_i\}_{i \in I}$ is any indexed collection of sets, we can form the Cartesian product

$$\prod_{i \in I} S_i.$$

The elements of the Cartesian product are "I-tuples" $\{s_i\}_{i \in I}$ such that each $s_i \in S_i$. For example, if the index set I is the set of natural numbers, and each of the S_i is the set of real numbers, then the Cartesian product is the set of *sequences of real numbers*.
Formally, an I-tuple is a function from I to $\bigcup_{i \in I} S_i$ such that

$$f(i) \in S_i, \ \forall i \in I \quad \text{and} \quad \prod_{i \in I} S_i = \{f : I \to \bigcup_{i \in I} S_i : f(i) \in S_i, \ \forall i \in I\}.$$

Example: If the index set I is the set of natural numbers $\mathbb{N}$, and each of the S_i is the set of real numbers, what is the Cartesian product $\prod_{i \in I} S_i$.?

Axiom of Choice (AC)

A Cartesian product of non-empty sets is non-empty.

Unfortunately, while the ZF axioms allow the construction of arbitrary Cartesian products, they don't guarantee that this product is non-empty, even if all of the original sets are non-empty. In the example above, where we take a Cartesian product of infinitely many copies of the real numbers, indexed by the natural numbers, we can see directly that this Cartesian product is non-empty. This is because we can explicitly specify an element, for example the constant sequence

$$(0, 0, 0, ...).$$

But in analysis, we often make arguments that involve choosing infinite sequences from arbitrary unknown (nonempty) sets. Such arguments are not allowed in ZF set theory.

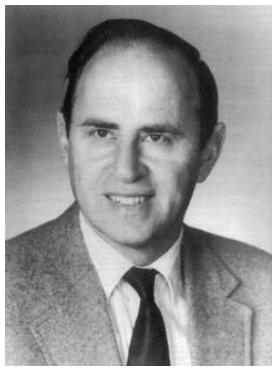
Challenge: Describe a rule to select a number from an arbitrary non-empty set of real numbers.

The solution is to add another axiom to ZF, called the *Axiom of Choice (AC)*, which asserts that a Cartesian product of non-empty sets is always non-empty (and thus one can "choose" an element of this Cartesian product).

ZFC

The Axiom of Choice has many equivalent formulations (many of which sound very different from the statement described above). It was eventually shown (in the 1960s) that AC is *logically independent* of ZF set theory. This means that it can neither be proven nor disproven from the other axioms.

Historically,
the use of AC was controversial, and there was some effort towards formulating parts of mathematics without appealing to AC. However, AC is now generally accepted and used in mathematics. Set theory based on the ZF axioms together with the Axiom of Choice is known as *ZFC*. In this course we will always assume AC, and generally not comment on its use (except maybe in parts of this chapter).



Kurt Godel and Paul Cohen proved the *logical independence* of the Axiom of Choice from ZF

Section 3

Functions

Once we have sets, we can consider *functions* between sets.

If A and B are sets, a function $f : A \to B$ is a "rule" which associates to each element of A exactly one element of B.

Formally, we can think of a function as a set of ordered pairs (x, y), where each x is an "input" and y is the corresponding "output". From this point of view, functions from A to B can be defined as subsets of the Cartesian product $A \times B$:

$f \subseteq A \times B$ is a function if and only if $\forall x \in A$, $\exists! y \in B$ such that $(x, y) \in f$.

In this course we will usually not be so formal.

Injective, surjective, and bijective functions

A function $f : A \to B$ is called:

- *injectve* (or "$1-1$") - every element of B is assigned to *at most* one element of A. Formally:

$$\forall x_1, x_2 \in A, f(x_1) = f(x_2) \implies x_1 = x_2.$$

  (Sometimes written $f : A \hookrightarrow B$.)
- *surjective* (or "onto") - every element of B is assigned to *at least* one element of A. Formally:

$$\forall y \in B, \exists x \in A \text{ such that } f(x) = y.$$

  (Sometimes written $f : A \twoheadrightarrow B$.)
- *bijective* - every element of B is assigned to *exactly* one element of A (=injective+surjective). Formally:

$$\forall y \in B, \exists! x \in A \text{ such thatt } f(x) = y.$$

Section 4

Comparison of sets

Question: Suppose you have two piles of coins. How can you tell which pile has more coins?

Some ideas:

- You can try to eyeball the piles and see which pile looks bigger, but this might be tricky if the coins are of many different sizes.
- If you're a sophisticated mathematician, you can simply count the number of coins in each pile, and check which number is larger.

However, this is roundabout and unnecessary - answering the question of which pile has more coins doesn't require the use of numbers. A straightforward approach is simply to match coins from the two piles against each other, one at a time. Whichever pile runs out first has fewer coins; or if they run out at the same time, then they have the same number of coins.

Definition

We say that two sets A and B have the same *cardinality* if there is a **bijection** $f : A \to B$; we then write $A \sim B$.

Note: Such a bijection is not unique (if the sets have more than one element)!

Example: The set of integers $\mathbb{Z}$ has the same cardinality as the set of even integers $2\mathbb{Z}$.

Example: The empty set $\emptyset$ does **not** have the same cardinality as the set $\{0\}$. More generally, the sets $\{0, 1, 2, ..., m\}$ and $\{0, 1, 2, ..., n\}$, with $m \neq n$, do not have the same cardinality ("pigeonhole principle").

Example: The set $\mathbb{N}$ has the same cardinality as the set $\mathbb{N} \times \mathbb{N}$.

Proof: Think of $\mathbb{N} \times \mathbb{N}$ as a grid:

$$\begin{array}{ccccc} (0,0) & (0,1) & (0,2) & (0,3) & \dots \\ (1,0) & (1,1) & (1,2) & (1,3) & \dots \\ (2,0) & (2,1) & (2,2) & (2,3) & \dots \\ (3,0) & (3,1) & (3,2) & (3,3) & \dots \\ \dots & \dots & \dots & \dots & \dots \end{array}$$

and consider the following labelling (follow the diagonals):

$$\begin{array}{cccc} (0,0)^{0} & (0,1)^{2} & (0,2)^{5} & (0,3)^{9} & \dots \\ (1,0)^{1} & (1,1)^{4} & (1,2)^{8} & (1,3)^{?} & \dots \\ (2,0)^{3} & (2,1)^{7} & (2,2)^{?} & (2,3)^{?} & \dots \\ (3,0)^{6} & (3,1)^{?} & (3,2)^{?} & (3,3)^{?} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{array}$$

Convince yourself this procedure determines a bijection between $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$. **Exercise:** Write down an explicit formula for the bijection.

Theorem (Cantor's Theorem)

Let S be any set, and let $\mathcal{P}(S)$ be its power set. Then $S \not\sim \mathcal{P}(S)$.

Example: Let S be any set, and let $\mathcal{P}(S)$ be its power set. Then $S \not\sim \mathcal{P}(S)$.

Proof: The argument is similar to Russell's Paradox. Suppose, on the contrary, that there is a bijection $f : S \to \mathcal{P}(S)$. Consider the set

$$T = \{x \in S : x \notin f(x)\}.$$

Since f is surjective, we have $T = f(y)$ for some y.
Question: Is $y \in T$?
If *yes*, then by definition of T, we have $y \notin f(y) = T$, which is a contradiction. If *no*, then we have $y \notin T = f(y)$, so by definition of T we have $y \in T$, which is again a contradiction.

For any sets A, B, and C, we have:

- $A \sim A$ (*reflexive*)
- $A \sim B \implies B \sim A$ (*symmetric*)
- $A \sim B \ \& \ B \sim C \implies A \sim C$ (*transitive*)

A relation satisfying these three conditions is called an *equivalence relation*. An equivalence relation on a set partitions the set into disjoint equivalence classes. We might then intuitively think of a "*cardinal number*" as an equivalence class of sets with the same cardinality. For example, we might think of the cardinal number "1" as the equivalence class of all sets with exactly one element. (Note that this does not make precise sense since there is no "set of all sets" to apply the equivalence relation to.) In any case, we write $|A| = |B|$ if $A \sim B$, and refer to $|A|$ as the "*cardinality*" of A.

We use the following notation:

1. $|A| = |B|$ if $A \sim B$
2. $|A| \leq |B|$ if there is an *injective* function $f : A \to B$.
3. $|A| < |B|$ if $|A| \leq |B|$ and $|A| \neq |B|$.

We refer to $|A|$ as the "*cardinality*" of A.

Note: In the course notes, the definition of $|A| \leq |B|$ is that there is a *surjective* function $f : B \to A$. Convince yourself that (assuming AC) these definitions are equivalent when A is non-empty.

For any sets A, B, and C, we have:

- $|A| \leq |A|$ (*reflexive*)
- $|A| \leq |B|$ & $|B| \leq |A| \implies |A| = |B|$ (*anti-symmetric*)
- $|A| \leq |B|$ & $|B| \leq |C| \implies |A| \leq |C|$ (*transitive*)

Theorem (Schroeder-Bernstein)

Let A and B be sets, and suppose that there exist injective functions $f : A \to B$ and $g : B \to A$. Then there exists a bijective function $h : A \to B$.

Ernst Schroeder and Felix Bernstein

Exercise

Show that $[0, 1]$ and $[0, 1)$ have the same cardinality.

Exercise

Give another proof that $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$ have the same cardinality, using the Schroeder-Bernstein Theorem.

Section 5

Finiteness

We use the positive integer n to denote the cardinality of the set $\{1, ..., n\}$, and 0 for the cardinality of the empty set.

Definition

A set S is *finite* $|S| = \{1, ..., n\}$, for some $n \in \mathbb{N}$. Otherwise S is *infinite*.

Exercise: Show that S is infinite if and only if $|S| \geq |\mathbb{N}|$.

Proof: We can see from the definition of finiteness that if S is finite, then $|S| < |\mathbb{N}|$.
So suppose S is infinite. Since S is infinite, S is non-empty. Choose $x_0 \in A$. Then $S \backslash \{x_0\}$ is non-empty (since otherwise $S = \{x_0\}$ would be finite). Next, choose $x_1 \in S \backslash \{x_0\}$. Continuing this way, we can find a sequence of distinct points in S indexed by $\mathbb{N}$, i.e. an injection of $\mathbb{N}$ into S. (This last step is subtle, and requires AC!)

A more conceptual definition is the notion of *Dedekind-finiteness*.

Definition

A set S is *Dedekind-infinite* if there is a bijection from S to a **proper** subset of itself. Otherwise S is Dedekind-finite.

Example: The set $\mathbb{N}$ is Dedekind-infinite.

It is intuitively clear that any finite set is Dedekind-finite (since the set $\{1, ..., n\}$ cannot be put in bijection with a proper subset).
Conversely, since any infinite set S satisfies $|S| \geq |\mathbb{N}|$, we see that any infinite set is Dedekind-infinite (why?). So the two notions are equivalent (assuming AC).

Exercise: Finiteness and Dedekind finiteness are equivalent (assuming AC)

Section 6

Countability

Comparison with $\mathbb{N}$

We have seen that a S is infinite if and only if $|S| \geq |N|$. So given a set S, there are three possibilities:

- $|S| < |\mathbb{N}|$ (i.e. S is finite)
- $|S| = |\mathbb{N}|$
- $|S| > |\mathbb{N}|$

Definition

We say that a set S is *countable* if $|S| \leq |\mathbb{N}|$. Otherwise S is *uncountable*. If S is countable and infinite we say that S is *countably infinite*.

Example: The set $\mathbb{Q}$ is countable.

Proof: Express each rational number as $\frac{a}{b}$ with $a, b \in \mathbb{Z}$, and send to $(a, b) \in \mathbb{Z} \times \mathbb{Z}$. This is an injection, and so we have

$$|\mathbb{Q}| \leq |\mathbb{Z} \times \mathbb{Z}| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$$

(where we have used the fact that $|\mathbb{Z}| = |\mathbb{N}|$).

Example: The set $\mathcal{P}(\mathbb{N})$ is uncountable.

Proof: We saw earlier that for any set S, we have $|S| \neq |\mathcal{P}(S)|$. On the other hand, we can inject S into $\mathcal{P}(S)$ (by $x \mapsto \{x\}$). So $|\mathcal{P}(S)| > |S|$, and therefore in particular $|\mathcal{P}(\mathbb{N})| > |\mathbb{N}|$. Therefore $\mathcal{P}(\mathbb{N})$ is uncountable.

The set $\mathcal{P}(\mathbb{N})$ can be identified with the set of *infinite bitstrings* (sequences of 0s and 1s)

Example: The set $[0,1]$ is uncountable.

Proof 1: First, note that we can identify elements of $\mathcal{P}(\mathbb{N})$ with infinite *bitstrings* (sequences of 0s and 1s) as follows. Let S be an element of $\mathcal{P}(\mathbb{N})$, i.e. a subset of $\mathbb{N}$.Then we assign to S the bitstring $x_0x_1x_2\ldots$, where

$$x_n = \begin{cases} 1 & n \in S \\ 0 & \text{otherwise} \end{cases}.$$

For example,

$$\{0,3,5,6,8,\ldots\} \mapsto 100101101\ldots.$$

Now assign to each bitstring the decimal number $0.x_0x_1$. This gives an injection from $\mathcal{P}(\mathbb{N})$ to $[0,1]$, so $|\mathcal{P}(\mathbb{N})| \leq |[0,1]|$. Since $\mathcal{P}(\mathbb{N})$ is uncountable, so is $[0,1]$.

Proof 2: (*Cantor's diagonal argument*). Suppose, on the contrary, that $[0,1]$ is countable (so $|[0,1]| \leq |\mathbb{N}|$). Let $f : [0,1] \to \mathbb{N}$ be a surjection. We can think of f as giving an infinite list (indexed by $\mathbb{N}$) which contains all of the numbers in $[0,1]$. Let's try writing this list in decimal form. To simplify notation, let $f_n = f(n)$. Then we can write each f_n as a decimal expansion:

$$\begin{aligned} f_0 &= 0.f_{00}f_{01}f_{02}f_{03}... \\ f_1 &= 0.f_{10}f_{11}f_{12}f_{13}... \\ f_2 &= 0.f_{20}f_{21}f_{22}f_{23}... \\ f_3 &= 0.f_{30}f_{31}f_{32}f_{33}... \end{aligned},$$

where the numbers f_{ij} represent the decimal coefficients of the f_i. Let

$$x_k = \begin{cases} 1 & f_{kk} \neq 1 \\ 2 & f_{kk} = 1 \end{cases}.$$

Now consider the number given by the decimal expansion

$$x = 0.x_0x_1x_2x_3....$$

Then $x \in [0,1]$ but $x \neq f_n$ for any n (why?), which contradicts our assumption that f is surjective.

The following is a useful way to show that certain kinds of sets are countable.

Theorem

Let I be a countable set, and let $\{S_i\}_{i\in I}$ be a set of countable sets indexed by I. Then the union $\bigcup_{i\in I} S_i$ is countable.

Proof: Assume the hypothesis. Since I is countable, there is an injection $f : I \to \mathbb{N}$. Since each S_i is countable, we can choose an injection $f_i : S_i \to \mathbb{N}$ for each $i \in I$.We will now define an injection $g : \bigcup_{i\in I} S_i \to \mathbb{N} \times \mathbb{N}$. For each element $x \in \bigcup_{i\in I} S_i$, choose an $i \in I$ such that $x \in S_i$ (why is this possible?). Then let $g(x) = (f(i), f_i(x))$. Convince yourself that g is an in fact an injection from $\bigcup_{i\in I} S_i$ into $\mathbb{N} \times \mathbb{N}$! Since $\mathbb{N} \times \mathbb{N}$ is countable, so is $\bigcup_{i\in I} S_i$. (Note that we have used AC). **In words:**

A countable union of countable sets is countable.

Exercise: The set S of finite subsets of $\mathbb{N}$ is countable.

Section 7

Other properties of cardinality

Comparability of cardinalities

Given two sets A and B, is it always true that either $|A| \leq |B|$ or $|B| \leq |A|$?

As we have seen, the $\leq$ relation is reflexive, anti-symmetric, and transitive. Do we also have *comparability*? Given two sets A and B, is it necessarily true that at least one of $|A| \leq |B|$ or $|B| \leq |A|$ holds? The answer turns out to be yes, assuming AC. So $\leq$ actually satisfies the properties of a *total order* (though again, there isn't actually a "set of all cardinalities" to which this total order applies). One other interesting question is whether there is any cardinality strictly in between $\mathbb{N}$ and $\mathcal{P}(\mathbb{N})$. The *Continuum Hypothesis (CH)* states that there is not:

$$\text{If } |\mathbb{N}| \leq |A| \leq |\mathcal{P}(\mathbb{N})| \text{ then either } |A| = |\mathbb{N}| \text{ or } |A| = |\mathcal{P}(\mathbb{N})|$$

The *General Continuum Hypothesis (GHC)* states that this holds for any infinite set S:

If $|S| \leq |A| \leq |\mathcal{P}(S)|$ and S is infinite, then either $|A| = |S|$or $|A| = |\mathcal{P}(S)|$